

CONSENSUS WITH TERNARY MESSAGES

ALEX OLSHEVSKY*

Abstract. We provide a protocol for real-valued average consensus by networks of agents which exchange only a single message from the ternary alphabet $\{-1, 0, 1\}$ between neighbors at each step. Our protocol works on time-varying undirected graphs subject to a connectivity condition, has a worst-case convergence time which is polynomial in the number of agents and the initial values, and requires no knowledge about either the graph topologies or the number of agents for implementation.

Key words. consensus protocols, multi-agent systems, distributed control.

AMS subject classifications. 93A14, 93C55, 68Q85

1. Introduction. The average consensus problem asks for a protocol by means of which n agents with initial values $x_1(0), \dots, x_n(0)$ can compute the average $\frac{1}{n} \sum_{i=1}^n x_i(0)$ subject to time-varying restrictions on inter-agent communication. Recent years have seen a surge of interest in consensus protocols due to their widespread use as building blocks for distributed control laws in multi-agent systems; for example, consensus protocols have been used for formation maintenance [14, 10, 7], coverage control [18, 19, 3], network clock synchronization [17, 5, 23], distributed task assignment and partitioning [8], statistical inference in sensor networks [1, 25, 9], and many other contexts in which centralized control is absent and agent motion and varying interference can lead to repeated failures of communication or sensing.

Consensus protocols typically involve each node i maintaining a variable $x_i(t)$ which is updated from time $t-1$ to time t by setting $x_i(t)$ to be a convex combination of those $x_j(t-1)$ for which j and i are neighbors in some undirected graph $G(t) = (\{1, \dots, n\}, E(t))$. The graph sequence $G(t)$ is meant to capture the constraints that dictate which pairs of agents can communicate or sense each other at each time. It is assumed that each $G(t)$ has a self-loop at every node, so that $x_i(t)$ always depends on its previous value $x_i(t-1)$. A typical update is the Metropolis iteration (introduced within the context of consensus in [25]), defined as

$$x_i(t) = x_i(t-1) + \sum_{j \in N_i(t)} \frac{x_j(t-1) - x_i(t-1)}{\max(d_i(t), d_j(t))} \quad (1.1)$$

where $N_i(t)$ is the set of neighbors of node i in $G(t)$ and $d_i(t)$ is the degree of node i in $G(t)$. It is assumed that the graph sequence $G(t)$ satisfies the long-term connectivity condition

$$\text{For each } k, \text{ the graph } (\{1, \dots, n\}, \cup_{t=k}^{\infty} E(t)) \text{ is connected,} \quad (1.2)$$

meaning that the communication restrictions faced by the nodes do not disconnect the network into noncommunicating groups after some finite time k . Under this assumption, the Metropolis iteration of Eq. (1.1) has the property that

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{n} \sum_{i=1}^n x_i(t) \quad \text{for each } i = 1, \dots, n$$

*Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA (aolshev2@illinois.edu).

meaning that all agents succeed in converging to the average. We refer the reader to the papers [2, 12, 20, 22, 24, 25] for proofs of this and similar assertions.

Consensus protocols have proven to be useful for multi-agent control due to their attractive robustness properties (namely, that the communication sequence $G(t)$ can vary arbitrarily subject to the relatively weak long-term connectivity constraint of Eq. (1.2)) and their local, distributed nature (i.e., every agent updates based only on neighboring agents). A number of recent advances in multi-agent control have proceeded by reduction to an appropriately defined consensus problem [3, 10, 7, 17, 23]. Moreover, we note that under slightly stronger assumptions on the sequence $G(t)$, it is possible to design consensus algorithms with convergence time bounds which are polynomial in the number of agents n [21].

However, a limitation of consensus protocols lies in the assumption that agent i can update $x_i(t)$ as a function of the states $x_j(t-1)$ of its neighbors in the graph $G(t)$. While this is very natural in some settings (such as formation or coverage control, where $x_j(t)$ represents the position of agent j , which neighboring agent i can be reasonably assumed to sense) in other contexts (such as clock synchronization or statistical inference in sensor networks) this implies that the agents can exchange real numbers at every step, a clearly unrealistic assumption.

The focus of this paper is to rectify this by proposing a consensus protocol in which the nodes exchange a finite number of bits with their neighbors at each time t . Specifically, we provide a deterministic protocol for the nodes to exchange messages with neighbors and update $x_i(t)$ which has the following properties:

1. For every time-varying sequence $G(t)$ subject to a certain connectivity assumption, each $x_i(t)$ converges to the average of the initial values $\frac{1}{n} \sum_{i=1}^n x_i(0)$.
2. The number of bits that is transmitted from agent to agent at each time t is bounded above independently of t, n and all other problem parameters (in fact in our protocol the agents transmit a single element from the ternary alphabet $\{-1, 0, 1\}$ to each neighbor at every t).
3. The protocol does not require knowledge of either the graphs $G(t)$ or even of the number of nodes n .
4. The protocol has a worst-case convergence time which is provably polynomial in the number of nodes n and the initial values $x_i(0)$.

A number of consensus protocols in which agents exchange finitely many bits at every step have recently been proposed [4, 6, 11, 13, 21, 15, 16]. However, each of these protocols lacks at least two of the features (1)-(4) above. We remark that features (1) and (3) are particularly crucial since much of the attractiveness of consensus protocols arises from their abilities to cope with link failures and their local, distributed nature. Feature (4) is clearly useful: the appeal of any protocol is increased by the availability of polynomial worst-case convergence bounds. Finally, feature (2) ensures that we do not implicitly assume that an arbitrarily large amount of data can be transmitted on each link before the graph changes.

We begin by stating our main result in Section 2, where we provide an informal description as well as a formal statement of our protocol. The proof of the main convergence theorem is given in the subsequent Section 3. Some simulations of our protocol are provided in Section 4. Finally, Section 5 provides some conclusions and lists some open problems. We note in particular that our new consensus protocol makes more stringent than usual assumptions on the graph sequence $G(t)$ and has more storage than the ordinary consensus algorithms, and we pose the improvement of these features as open problems.

2. Our results. We first reprise the notation we have introduced: $G(t)$ is a sequence of undirected graphs with a self-loop at every node, $N_i(t)$ is the set of neighbors of node i in $G(t)$, and $d_i(t)$ is the degree of node i in $G(t)$.

2.1. Intuitive description of the protocol. We begin by informally describing the idea behind our protocol. We would like to run the Metropolis update of Eq. (1.1), but without the ability to transmit real numbers, node i will not know the values of its neighbors $x_j(t-1)$ exactly. Consequently, node i will maintain estimates of $x_j(t)$; we will use the notation $\hat{x}_{i,j,\text{in}}(t)$ for the estimate that node i has for $x_j(t)$.

A crucial point is that node i will be aware of the value of any neighbor j 's estimate of $x_i(t)$ because this estimate will be built from the messages node i has previously sent j . Thus node i will maintain the variable $\hat{x}_{i,j,\text{out}}(t)$ which will equal node j 's estimate of $x_i(t)$, i.e., will equal $\hat{x}_{j,i,\text{in}}(t)$.

At each time t , node i will receive a message from j from the alphabet $\{-1, 0, 1\}$. If it receives a $+1$, it will add $1/t^\alpha$ to $\hat{x}_{i,j,\text{in}}$; if it receives a -1 , it adds a $-1/t^\alpha$ to the same; and if it receives a zero, it leaves $\hat{x}_{i,j,\text{in}}$ unchanged. Naturally, node j decides to send a $1, 0, -1$ depending on whether $\hat{x}_{i,j,\text{in}}$ is too low by at least $1/t^\alpha$, within $1/t^\alpha$ of the true value, or too high by a factor of at least $1/t^\alpha$, respectively. Note that even though it is node i which maintains the variable $\hat{x}_{i,j,\text{in}}$, node j knows what it is by the previous paragraph.

The number α will be in $(0, 1)$ so that

$$\sum_{t=1}^n \frac{1}{t^\alpha} = +\infty$$

which will be key in ensuring that the estimates $\hat{x}_{i,j,\text{in}}$ increasingly become accurate for links that regularly appear.

Meanwhile, the nodes will implement the Metropolis update, with each node i replacing the neighboring values $x_j(t-1)$ by the estimates $\hat{x}_{i,j,\text{in}}$. However, node i will not include all its neighbors in the Metropolis updates; rather, it will include only those which sent it a zero at time t (which means that $\hat{x}_{i,j,\text{in}}$ will not be too far away from $x_j(t)$) and whose $\hat{x}_{i,j,\text{in}}$ are sufficiently far from $x_i(t)$ so that the inevitable perturbation error inherent in using imprecise estimates does not affect the available bounds the available L^2 Lyapunov bounds for the Metropolis algorithm.

Moreover, we introduce a stepsize of $1/t^\beta$ into the Metropolis algorithm where $\beta > \alpha$ and $\beta \in (0, 1)$. Indeed, by choosing $\beta > \alpha$, we ensure that agents change their values $x_j(t)$ slower than estimates $\hat{x}_{i,j,\text{in}}$ change. Intuitively, while the estimates $\hat{x}_{i,j,\text{in}}$ will get accurate by an additive factor of $1/t^\alpha$ whenever they are very inaccurate, the values $x_i(t)$ will have their movement attenuated by a factor of $1/t^\beta$. The introduction of the $1/t^\beta$ stepsize ensures that not only do the estimates “catch up” to the true values, but we can give a simple bound on how long it takes until estimates become accurate.

2.2. Formal description of the protocol. The nodes begin with initial values $x_i(0)$. Every node i will maintain variables $\hat{x}_{i,j,\text{in}}(t), \hat{x}_{i,j,\text{out}}(t)$ for every node which has been its neighbor in some past $G(t)$. Node i initializes

$$\hat{x}_{i,j,\text{in}}(t-1) = \hat{x}_{i,j,\text{out}}(t-1) = 0$$

at the first time t when the edge (i, j) belongs to $E(t)$.

At each iteration $t = 1, 2, 3, \dots$, node i will send a value from the set $\{-1, 0, 1\}$ to each of its neighbors j in $G(t)$. The value i sends to j is

$$q_{i \rightarrow j}(t) = R[t^\alpha(x_i(t-1) - \hat{x}_{i,j,\text{out}}(t-1))]$$

where

$$R[x] = \begin{cases} 1 & \text{if } x > 1 \\ 0 & \text{if } -1 \leq x \leq 1 \\ -1 & \text{if } x < -1 \end{cases}$$

Note that node i may send different messages to different neighbors. After sending $q_{i \rightarrow j}(t)$ to each neighbor $j \in N_i(t)$ and receiving $q_{j \rightarrow i}(t)$ from the same, node i updates the values $\hat{x}_{i,j,\text{in}}(t), \hat{x}_{i,j,\text{out}}(t)$ as

$$\begin{aligned} \hat{x}_{i,j,\text{out}}(t) &= \hat{x}_{i,j,\text{out}}(t-1) + \frac{q_{i \rightarrow j}(t)}{t^\alpha} \\ \hat{x}_{i,j,\text{in}}(t) &= \hat{x}_{i,j,\text{in}}(t-1) + \frac{q_{j \rightarrow i}(t)}{t^\alpha} \end{aligned}$$

and then updates its value $x_i(t)$ as

$$x_i(t) = x_i(t-1) + \frac{1}{t^\beta} \sum_{j \in S(i,t)} \frac{\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{i,j,\text{out}}(t)}{4\max(d_i(t), d_j(t))} \quad (2.1)$$

Here $S(i, t)$ is the set of neighbors j of node i in $G(t)$ which satisfy

$$|\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{i,j,\text{out}}(t)| > \frac{4}{t^\alpha}$$

and

$$q_{j \rightarrow i}(t) = q_{i \rightarrow j}(t) = 0.$$

2.3. Main result. We now provide the main convergence theorem for our protocol. We begin with a few definitions needed to specify the class of graph sequences on which our protocol is guaranteed to work.

DEFINITION 2.1. *We will call the graph sequence $G(t)$ B -core-connected if there exists a set of edges $E_\infty \subset \cup_t E(t)$ such that the graph $(\{1, \dots, n\}, E_\infty)$ is connected; and*

$$E_\infty \subset \cup_{t=kB+1}^{(k+1)B} E(t)$$

for every k .

That is, a sequence is B -core-connected if there is a set of edges, forming a connected graph, each of which appears in every interval $[kB+1, (k+1)B]$. We will say that the edges in E_∞ are *core edges*. We note that this is a stronger assumption than the B -connectivity assumption typically made in the analysis of consensus algorithms. We will discuss this point in more detail shortly.

Our main result will show provide bounds until our consensus protocol reduces a certain measure of disagreement to a small value forever. We now define this measure (this is $V_2(t)$) in the next definition) as well as several related concepts.

DEFINITION 2.2.

$$\begin{aligned} M(x) &= \max_{i=1,\dots,n} x_i \\ m(x) &= \min_{i=1,\dots,n} x_i \\ W(x) &= M(x) - m(x) \\ V_2(x) &= \sqrt{\sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2} \end{aligned}$$

Note that will use the natural shorthands $M(t), m(t), W(t), V_2(t)$ for $M(x(t)), m(x(t)), W(x(t)), V_2(x(t))$.

We can now state the main result of this paper.

THEOREM 2.3. *If $0 < \alpha < \beta < 1$ then for all nodes i , initial values $x(0)$, and B -core-connected sequences $G(t)$, it is true that*

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{n} \sum_{j=1}^n x_j(0).$$

Moreover, if¹

$$\begin{aligned} t \geq 2^{\frac{1}{1-\beta}} & \left(2^{\frac{2}{1-\alpha}} [4(1+B)W(0) + 2B]^{\frac{1}{\beta-\alpha}} + (20B\|x(0)\|_\infty)^{\frac{2}{1-\alpha}} + 26B + (128n^4 B)^{\frac{1}{1-\beta}} \right) \\ & + \max \left(\left(2^{\frac{1}{1-\beta}} 128n^4 B \log \frac{\epsilon}{V_2(0)} \right)^{\frac{1}{1-\beta}}, \left(\frac{8n^{1.5}}{\epsilon} \right)^{\frac{1}{\alpha}} \right) \end{aligned}$$

then

$$V_2(x(t)) \leq \epsilon$$

for all i, j .

Our result states that the consensus protocol of the previous section succeeds in computing the average on B -core-connected sequence and provides a bound for its convergence time.

REMARK 2.4. *While the above convergence time expression is somewhat unwieldy, we emphasize that it is polynomial in $n, \epsilon, B, W(0), V(0), \|x(0)\|_\infty$ for any choice of α and β satisfying the assumption $0 < \alpha < \beta < 1$.*

REMARK 2.5. *Observe that there is a tradeoff between the steady-state and transient terms in the convergence time. In particular, as $\epsilon \rightarrow 0$, the term $(8n^{1.5}/\epsilon)^{1/\alpha}$*

¹ $\lceil x \rceil$ means the smallest integer which is at least x .

is going to dominate all the other terms in the convergence time expression. Consequently, to obtain the best asymptotic decay rate we should choose α close to 1 which means we must also choose β close to 1. In that case, error decay at time t decays nearly as well as $O(n^{1.5}/t)$ as t approaches infinity. However, choosing $\alpha_n \rightarrow 1$ and $\beta_n \rightarrow 1$ causes the transient term to blow up, so that it takes longer and longer until the asymptotically dominant term dominates the other terms.

In general, there is no single best choice of α, β ; rather every choice gives us a tradeoff between transient bounds and steady-state decay. For example, choosing $\alpha = 3/4, \beta = 7/8$ gives us that the time until $V_2(x(t)) \leq \epsilon$ is

$$O\left((B + W(0) + BW(0))^8\right) + O(B\|x(0)\|_\infty)^8 + O(n^{32}B^8) + \max\left(O(n^{32}B^8(\log \frac{V_2(0)}{\epsilon})^8, O\left(\frac{n^2}{\epsilon^{4/3}}\right)\right)$$

Note that every term which does not have an ϵ in it will become negligible as $\epsilon \rightarrow 0$. In this limit, the dominant term will be $O(\frac{8n^2}{\epsilon^{4/3}})$ which will grow faster as $\epsilon \rightarrow 0$ than the logarithmic term $O(n^{32}B^8 \log \frac{V_2(0)}{\epsilon})$.

On the other hand, choosing $\alpha = 1/4, \beta = 1/2$ gives us that the time until $V_2(x(t)) \leq \epsilon$ is

$$O\left((B + W(0) + BW(0))^4\right) + O(B\|x(0)\|_\infty)^{8/3} + O(n^8B^2) + \max\left(O(n^8B^2(\log \frac{V_2(0)}{\epsilon})^2, O\left(\frac{n^6}{\epsilon^4}\right)\right)$$

which has a worse asymptotic term $O(\frac{n^6}{\epsilon^4})$ but the rest of the terms are smaller if $B, W(0), \|x(0)\|_\infty, V_2(0)$ are large.

REMARK 2.6. The analysis of consensus algorithms usually relies on a weaker notion of connectivity, namely the so-called B -connectivity (sometimes called uniform connectivity) condition: a sequence $G(t)$ of undirected graphs is called B -connected if for each k , the graph

$$\left(\{1, \dots, n\}, \cup_{t=kB+1}^{(k+1)B} E(t)\right)$$

is connected. This encompasses a wider class of sequences compared to notion of B -core-connectivity which we use, which requires not only the above graph to be connected, but also that it has the same connected subgraph for every k .

However, it is easy to see that our convergence proof works on B -connected sequences rather than just B -core connected sequences; thus we have that the convergence result $\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{n} \sum_{j=1}^n x_j(0)$ holds for all B -connected sequence. We will justify this point in a remark following the main proof. However, the convergence-time result is only known to hold for B -core-connected sequences, and it is an open problem whether one can obtain a similar convergence time on B -connected sequences.

REMARK 2.7. Our protocol requires node i to keep track of the numbers $\hat{x}_{i,j,\text{in}}, \hat{x}_{i,j,\text{out}}$ for every other node j it interacts with. By contrast, the standard consensus algorithm keeps track of only a single real number, namely $x_i(t)$ at node i . Unfortunately, it seems that a “storage blowup” phenomenon of this sort is unavoidable, though it remains an open question to prove this in a formal sense.

As a consequence, our consensus protocol is most attractive on graphs sequences $G(t)$ in which every node interacts with a relatively small number of neighbors. One such example is the geometric random graph model of wireless networks, in which

sensors have locations in $[0, 1]^2$ and every node is connected to $O(\log n)$ neighbors []. Our consensus protocol will work in such networks, even if unpredictable interference or malicious jamming makes links unreliable, with every node needing to store only $O(\log n)$ additional estimates.

3. Main proof. This section is devoted to the proof of our main result, Theorem 2.3. We will begin with a long sequence of lemmas which will establish properties of our consensus protocol, and put off the calculations that lead to the bound of Theorem 2.3 until the last possible step.

Our first lemma states the natural fact that the estimate $\hat{x}_{j,i,\text{out}}$ maintained by node j equals the estimate $\hat{x}_{i,j,\text{in}}$ maintained by node i . Intuitively, this means that node j knows node i 's estimate of its value $x_j(t)$.

LEMMA 3.1.

$$\hat{x}_{i,j,\text{in}}(t) = \hat{x}_{j,i,\text{out}}(t)$$

whenever these variables are defined.

REMARK 3.2. Recall that the variables $\hat{x}_{i,j,\text{in}}(t)$ and $\hat{x}_{j,i,\text{out}}(t)$ are initialized and maintained at the first time (i, j) appears in the sequence $G(t)$.

Proof. The first time $(i, j) \in E(t)$, the variables $\hat{x}_{i,j,\text{in}}(t-1)$ and $\hat{x}_{j,i,\text{out}}(t-1)$ are initialized to zero by both nodes i and j . Subsequently, they are updated as (see Section 2.2)

$$\begin{aligned}\hat{x}_{j,i,\text{out}}(t) &= \hat{x}_{j,i,\text{out}}(t-1) + \frac{q_{j \rightarrow i}(t)}{t^\alpha} \\ \hat{x}_{i,j,\text{in}}(t) &= \hat{x}_{i,j,\text{in}}(t-1) + \frac{q_{j \rightarrow i}(t)}{t^\alpha}\end{aligned}$$

at each time t such that $(i, j) \in E(t)$. Consequently, they have the same value at all times. \square

The next corollary tells us that if i uses $\hat{x}_{i,j,\text{in}}(t)$ in its Metropolis update of Eq. (2.1), then j uses $\hat{x}_{j,i,\text{out}}(t)$ in its own update at that time as well. Thus our protocol in a sense maintains the natural symmetry of the graphs $G(t)$.

COROLLARY 3.3. $j \in S(i, t)$ if and only if $i \in S(j, t)$.

Proof. Suppose $j \in S(i, t)$. This happens if and only if $q_{i \rightarrow j}(t) = q_{j \rightarrow i}(t) = 0$ and

$$|\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{j,i,\text{out}}(t)| > \frac{4}{t^\alpha}. \quad (3.1)$$

The condition $q_{i \rightarrow j}(t) = q_{j \rightarrow i}(t) = 0$ is naturally symmetric between i and j . Moreover, using Lemma 3.1, we have that Eq. (3.1) is equivalent to

$$|\hat{x}_{j,i,\text{out}}(t) - \hat{x}_{i,j,\text{in}}(t)| > \frac{4}{t^\alpha},$$

so that $i \in S(j, t)$. \square

We next remark that the update equations of our protocols may be rewritten in more convenient matrix form.

REMARK 3.4. For $i \neq j$, define

$$w_{ij}(t-1) = \begin{cases} 0 & \text{if } j \notin S(i, t) \\ \frac{\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{i,j,\text{out}}(t)}{x_j(t-1) - x_i(t-1)} & \text{else} \end{cases}$$

Then

$$x_i(t) = x_i(t-1) + \frac{1}{t^\beta} \sum_{j \in S(i, t)} \frac{w_{ij}(t-1)}{4\max(d_i(t), d_j(t))} (x_j(t-1) - x_i(t-1)). \quad (3.2)$$

We will subsequently require some bounds on the sizes of the weights $w_{ij}(t-1)$ from the above equation. These are provided in the following lemma.

LEMMA 3.5. If $w_{ij}(t-1) \neq 0$ then

$$\frac{2}{3} \leq w_{ij}(t-1) \leq 2$$

Proof. Indeed, since $w_{ij}(t-1) \neq 0$ we have that $j \in S(i, t)$ which means

$$|\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{i,j,\text{out}}(t)| > \frac{4}{t^\alpha} \quad (3.3)$$

Moreover, we also have that $q_{i \rightarrow j}(t) = 0$, which means that

$$|\hat{x}_{i,j,\text{out}}(t) - x_i(t-1)| = |\hat{x}_{i,j,\text{out}}(t-1) - x_i(t-1)| \leq \frac{1}{t^\alpha}$$

and similarly, $q_{j \rightarrow i}(t) = 0$ which means

$$|\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)| = |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-1)| = |\hat{x}_{j,i,\text{out}}(t-1) - x_j(t-1)| \leq \frac{1}{t^\alpha}.$$

where the second equality used Lemma 3.1. We therefore have

$$|\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{i,j,\text{out}}(t)| - |x_j(t-1) - x_i(t-1)| \leq \frac{2}{t^\alpha}$$

and combining this with Eq. (3.3),

$$w_{ij}(t-1) = \frac{\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{i,j,\text{out}}(t)}{x_j(t-1) - x_i(t-1)} \in \left[\frac{4}{6}, \frac{4}{2}\right],$$

which completes the proof. \square

REMARK 3.6. As a consequence of Lemma 3.1, we have that $w_{ij}(t) = w_{ji}(t)$ for all i, j, t .

Our next lemma rewrites the update dynamics of our consensus protocol in still more convenient form.

LEMMA 3.7. *We may rewrite the main update Eq. (2.1) as*

$$x(t) = \left(1 - \frac{1}{t^\beta}\right)x(t-1) + \frac{1}{t^\beta}A(t-1)x(t-1),$$

where $A(t-1)$ is a symmetric, stochastic, diagonally dominant matrix with the property that if $a_{ij}(t-1)$ is positive, then $a_{ij}(t-1) \geq 1/(8 \max(d_i(t), d_j(t)))$.

Proof. From Eq. (3.2),

$$\begin{aligned} x_i(t) &= \left(1 - \frac{1}{t^\beta} \sum_{j \in S(i,t)} \frac{w_{ij}(t-1)}{4 \max(d_i(t), d_j(t))}\right) x_i(t-1) + \frac{1}{t^\beta} \sum_{j \in S(i,t)} \frac{w_{ij}(t-1)}{4 \max(d_i(t), d_j(t))} x_j(t-1) \\ &= \left(1 - \frac{1}{t^\beta}\right) x_i(t-1) + \frac{1}{t^\beta} \left(1 - \sum_{j \in S(i,t)} \frac{w_{ij}(t-1)}{4 \max(d_i(t), d_j(t))}\right) x_i(t-1) + \frac{1}{t^\beta} \sum_{j \in S(i,t)} \frac{w_{ij}(t-1)}{4 \max(d_i(t), d_j(t))} x_j(t-1) \end{aligned}$$

We therefore define $A(t-1)$ in the natural way from the above equation by setting

$$a_{ij} = \frac{w_{ij}(t-1)}{4 \max(d_i(t), d_j(t))}$$

for all $i \neq j$ and

$$a_{ii} = 1 - \sum_{j \in S(i,t)} \frac{w_{ij}(t-1)}{4 \max(d_i(t), d_j(t))}$$

By construction the rows add up to 1. The off-diagonal entries are clearly nonnegative, and the diagonal entries are not just diagonal but at least $1/2$ because Lemma 3.5 implies that

$$1 - \sum_{j \in S(i,t)} \frac{w_{ij}(t)}{4 \max(d_i(t), d_j(t))} \geq \frac{1}{2}$$

This makes $A(t)$ nonnegative, stochastic, and diagonally dominant. Finally, Remark 3.6 implies that $A(t-1)$ is symmetric. \square

The next lemma lists some natural Lyapunov functions for our protocol.

LEMMA 3.8. *$M(t)$ is nonincreasing, $m(t)$ is nondecreasing, and $V_2(x(t))$ is nonincreasing.*

Proof. The first two claims follow immediately from the stochasticity of $A(t-1)$. The last one follows from the symmetry of A - see [21] for a proof. \square

The following lemma establishes that the sizes of the estimates possessed by the nodes cannot be too large.

LEMMA 3.9. *For all i, j, t such that $\hat{x}_{i,j,\text{in}}(t)$ is defined,*

$$|\hat{x}_{i,j,\text{in}}(t)| \leq \|x(0)\|_\infty.$$

Proof. By induction, $\hat{x}_{i,j,\text{in}}(t)$ belongs to the convex hull of 0 and $x_j(0), x_j(1), x_j(2), \dots, x_j(t)$. Consequently, its absolute value cannot be larger than $\max(|x_j(0)|, |x_j(1)|, \dots, |x_j(t)|)$. By Lemma 3.8, this is at most $\|x(0)\|_\infty$. \square

We now proceed to our first substantial lemma, which proves that after some finite number of steps which depends polynomially on the initial condition $x(0)$, the estimates $\hat{x}_{i,j,\text{in}}(t)$ associated with core links (i, j) are close to the correct values $x_j(t-1)$.

LEMMA 3.10. *If*

$$t \geq 2^{\frac{2}{1-\alpha}} \lceil 8W(0) + 3B \rceil^{\frac{1}{\beta-\alpha}} + (20B\|x(0)\|_\infty)^{\frac{2}{1-\alpha}} + 20B$$

then

$$|\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)| \leq \frac{1}{(t-2B-1)^\alpha}$$

for every pair of nodes i, j such that the edge (i, j) is a core edge.

Proof. First note that as a consequence of Lemma 3.8 and 3.7, we can easily bound how much nodes change value from one step to the next. Indeed, Lemma 3.8 implies that for any t ,

$$\max_{k,l} |x_k(t) - x_l(t)| \leq W(0),$$

and therefore from Eq. (3.2),

$$|x_j(t) - x_j(t-1)| \leq \frac{(1/2)W(0)}{t^\beta}.$$

We now use this observation to relate the message $q_{j \rightarrow i}(t)$ to the estimation error $|\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)|$. Indeed, observe that $q_{j \rightarrow i}(t) = 0$, then

$$|\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)| = |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-1)| \leq \frac{1}{t^\alpha}.$$

On the other hand, if $q_{j \rightarrow i}(t) \in \{-1, 1\}$, then

$$\begin{aligned} |\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)| &\leq |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-1)| - \frac{1}{t^\alpha} \\ &\leq |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-2)| + |x_j(t-1) - x_j(t-2)| - \frac{1}{t^\alpha} \\ &\leq |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-2)| + \frac{(1/2)W(0)}{(t-1)^\beta} - \frac{1}{t^\alpha} \end{aligned}$$

Now if t is large enough so that $t > 1 + (8W(0))^{\frac{1}{\beta-\alpha}}$, we will then have the upper bound $W(0) \leq (1/8)(t-1)^{\beta-\alpha}$ which allows us to eliminate $W(0)$ from the above equation: we get that if $q_{j \rightarrow i}(t) \in \{-1, 1\}$,

$$|\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)| \leq |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-2)| + \frac{1/16}{(t-1)^\alpha} - \frac{1}{t^\alpha} \leq |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-2)| - \frac{7/8}{t^\alpha} \quad (3.4)$$

where we needed that $t > 1$ for the final inequality. Thus, for large enough t , the estimation error $|\widehat{x}_{i,j,\text{in}}(t) - x_j(t-1)|$ either decreases at each step or is already small. We now proceed to turn this into a quantitative bound on $|\widehat{x}_{i,j,\text{in}}(t) - x_j(t-1)|$.

Let Λ be the set of times when the link (i, j) appears in the graph sequence. Suppose $z > 1 + (8W(0))^{\frac{1}{\beta-\alpha}}$ and $t \geq z$; by considering the last element in Λ in $[z, t]$ we obtain the inequality

$$|\widehat{x}_{i,j,\text{in}}(t) - x_j(t-1)| \leq \max \left(2\|x(0)\| - \sum_{k \in \Lambda \cap [z, t]} \frac{7/8}{k^\alpha}, \max_{t'=z, \dots, t} \frac{1}{(t')^\alpha} - \sum_{k \in \Lambda \cap [t'+1, t]} \frac{7/8}{k^\alpha} \right) \quad (3.5)$$

Note that to justify this inequality we needed the claim that $|\widehat{x}_{i,j,\text{in}}(z-1) - x_j(z-2)| \leq 2\|x(0)\|_\infty$ which is justified by appealing to Lemmas 3.8 and 3.9.

Next we invoke the assumption that every core edge that appears in each block of times $[kB+1, (k+1)B]$. This means that because Λ has nonzero intersection with each $[kB+1, (k+1)B]$, we can infer that for any $[a, b]$ such that $a \geq 2B$ and $b-a \geq 2B$ we have²:

$$\sum_{k \in \Lambda \cap [a, b]} \frac{7/8}{k^\alpha} \geq \frac{1}{2B} \sum_{k=a}^b \frac{1/5}{k^\alpha}.$$

We now plug this lower bound into Eq. (3.5), while simultaneously setting $z = \lceil (8W(0) + 3B)^{\frac{1}{\beta-\alpha}} \rceil$ (which is of course always at least $2B$, strictly larger than 1, and strictly larger than $1 + (8W(0))^{\beta-\alpha}$). Note that the lower bound on t assumed in the statement of this lemma certainly implies $t - (z+1) \geq 2B$; thus we have

$$|\widehat{x}_{i,j,\text{in}}(t) - x_j(t-1)| \leq \max \left(2\|x(0)\|_\infty - \sum_{k=\lceil 8W(0)+3B \rceil^{\frac{1}{\beta-\alpha}}}^t \frac{1/10}{Bk^\alpha}, \max_{t'=\lceil (8W(0)+3B)^{\frac{1}{\beta-\alpha}} \rceil, \dots, t-2B-1} \frac{1}{t'^\alpha} - \sum_{k=t'+1}^t \frac{1/10}{Bk^\alpha}, \frac{1}{(t-2B)^\alpha} \right) \quad (3.6)$$

Our next step is to lower bound the sums appearing on the right-hand side of the above equation. Using the standard method of lower-bounding the sum of a nonin-

²This inequality follows by some elementary manipulations, which we spell out in this footnote so as to avoid interrupting the flow of the above proof. Indeed, note that every interval of length $2B$ or more has at least one element from Λ . Moreover, since $1/k^\alpha$ is a decreasing function of k , we have that

$$\sum_{k \in \Lambda \cap [a, b]} \frac{1}{k^\alpha} \geq \sum_{k \geq 1 \text{ such that } a+k2B \leq b} \frac{1}{(a+k2B)^\alpha}$$

Now since $a + 2B \leq 2a$ (due to the assumption $a \geq 2B$), we have that $1/(a+2B)^\alpha \geq 1/(2a)^\alpha$ so that

$$\sum_{k \in \Lambda \cap [a, b]} \frac{1}{k^\alpha} \geq \left(\frac{1}{2} \frac{1}{(a+2B)^\alpha} + \frac{1}{2} \frac{1}{(a+2B)^\alpha} + \sum_{k \geq 2 \text{ such that } a+k2B \leq b} \frac{1}{2k^\alpha} \right) \geq \sum_{k \geq 0 \text{ such that } a+k2B \leq b} \frac{1}{2 \cdot 2^\alpha k^\alpha}$$

which implies

$$\sum_{k \in \Lambda \cap [a, b]} \frac{1}{k^\alpha} \geq \frac{1}{2B} \sum_{k=a}^b \frac{1/4}{k^\alpha}.$$

creasing function by an integral, we have

$$\sum_{k=a}^b \frac{1}{k^\alpha} \geq \frac{(b+1)^{1-\alpha} - a^{1-\alpha}}{1-\alpha} = \frac{((b+1)^{\frac{1-\alpha}{2}} + a^{\frac{1-\alpha}{2}})((b+1)^{\frac{1-\alpha}{2}} - a^{\frac{1-\alpha}{2}})}{1-\alpha}. \quad (3.7)$$

Now the lower bound on t we have assumed in the statement of this lemma implies that $t \geq 2^{\frac{2}{1-\alpha}} [8W(0) + 3B]^{\frac{1}{\beta-\alpha}}$ and $t^{(1-\alpha)/2} \geq 20B \|x(0)\|_\infty$ so that Eq. (3.7) implies that the first term in the maximization of Eq. (3.6) is negative. As for the second term in the maximization, it is therefore bounded as

$$\max_{t' = \lceil 8W(0) + 3B \rceil^4, \dots, t-2B-1} \frac{1}{(t')^\alpha} - \frac{1}{10B} \frac{((t+1)^{1-\alpha} - (t'+1)^{1-\alpha})}{1-\alpha}$$

and since $t' \geq 20B$, the expression inside the maximum is an increasing function³ of t' . Consequently, we may choose $t' = t - 2B - 1$ in the second term, which immediately implies that

$$|\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)| \leq \frac{1}{(t-2B-1)^\alpha}$$

□

We next demonstrate a consequence of the previous lemma: if t is large enough and the message $q_{i \rightarrow j}$ on a core edge is nonzero at some time, then the same message $q_{i \rightarrow j}$ will be zero for a substantial amount of time whenever the edge (i, j) appears again.

LEMMA 3.11. *Suppose*

$$t \geq 2^{\frac{2}{1-\alpha}} [8(1+B)W(0) + 3B]^{\frac{1}{\beta-\alpha}} + (20B \|x(0)\|_\infty)^{\frac{2}{1-\alpha}} + 20B + 1$$

and the edge (i, j) is a core edge. If $q_{i \rightarrow j}(t) \neq 0$ then $q_{i \rightarrow j}(k) = 0$ for all $k \in [t+1, t+3B]$ during which (i, j) appears.

Proof. By Lemma 3.10,

$$|\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-2)| \leq \frac{1}{(t-2B-2)^\alpha}. \quad (3.8)$$

Consequently, $q_{i \rightarrow j}(t) \neq 0$ implies that for each $k \in [t, t+3B-1]$

$$\begin{aligned} |\hat{x}_{i,j,\text{in}}(k) - x_j(k)| &\leq |\hat{x}_{i,j,\text{in}}(k) - x_j(k-1)| + \frac{(1/2)W(0)}{k^\beta} \\ &\leq |\hat{x}_{i,j,\text{in}}(t) - x_j(t-1)| + 3B \frac{(1/2)W(0)}{t^\beta} \\ &\leq |\hat{x}_{i,j,\text{in}}(t-1) - x_j(t-2)| - \frac{7/8}{t^\alpha} + 3B \frac{(1/2)W(0)}{t^\beta} \\ &\leq \frac{1}{(t-2B-2)^\alpha} - \frac{7/8}{t^\alpha} + 3B \frac{(1/2)W(0)}{t^\beta}. \end{aligned} \quad (3.9)$$

³Indeed, if $f(t) = c/t^\alpha + d(t+1)^{1-\alpha}$ where $\alpha \in (0, 1)$ then

$$f'(t) = -\alpha c t^{-\alpha-1} + d(1-\alpha)(t+1)^{-\alpha}.$$

The latter expression is nonnegative whenever $\frac{d(1-\alpha)}{c\alpha} \frac{t^{\alpha+1}}{(t+1)^\alpha} \geq 1$.

Here the first two inequalities follow from the observation that $x_j(t)$ and $|\widehat{x}_{i,j,\text{in}}(t) - x_j(t-1)|$ can increase by at most $\frac{(1/2)W(0)}{t^\beta}$ as t increments by one; the third follows from Eq. (3.4); and the last one follows by plugging in Eq. (3.8).

We claim the the assumption $t \geq 2\lceil 8(1+B)W(0) + 3B \rceil^{\frac{1}{\beta-\alpha}} + (20BV_\infty(0))^{\frac{2}{1-\alpha}} + 20B + 1$ implies the final expression on the right is always at most $1/(t+3B)^\alpha$. Indeed, this assumption implies that $t \geq 20B$ so that $t - 2B - 2 \geq 0.80t$ and $t + 3B \leq 1.15t$. Consequently,

$$\begin{aligned} \frac{1}{(t+3B)^\alpha} - \frac{1}{(t-2B-2)^\alpha} + \frac{7/8}{t^\alpha} &\geq \frac{1/1.15^\alpha - 1/0.8^\alpha + 7/8}{t^\alpha} \\ &\geq \frac{3/8}{t^\alpha} \\ &= \frac{3/8}{t^\beta} t^{\beta-\alpha} \\ &\geq \frac{3/2}{t^\beta} BW(0) \end{aligned}$$

where the last step used our assumed lower bound on t which implies that $t^{\beta-\alpha} \geq 4BW(0)$. The last inequality is, after rearranging, precisely our claim that the right-hand side of Eq. (3.9) is at most $1/(t+3B)^\alpha$.

Since

$$|\widehat{x}_{i,j,\text{in}}(k) - x_j(k)| \leq \frac{1}{(t+3B)^\alpha} \leq \frac{1}{k^\alpha}$$

for all $k \in [t, t+3B-1]$, we have that $q_{i \rightarrow j}(k) = 0$ for all $k \in [t+1, t+3B]$ when the edge (i, j) appears. \square

A corollary of this lemma is that every $3B$ steps, both messages $q_{i \rightarrow j}$ and $q_{j \rightarrow i}$ across a core link will be zero. We state this formally next.

COROLLARY 3.12. *Suppose the edge (i, j) is a core edge and suppose*

$$t \geq 2^{\frac{2}{1-\alpha}} \lceil 4(1+B)W(0) + 2B \rceil^{\frac{1}{\beta-\alpha}} + (20B\|x(0)\|_\infty)^{\frac{2}{1-\alpha}} + 20B.$$

Then there exists a time in $[t+1, t+3B]$ with

$$q_{i \rightarrow j}(t) = q_{j \rightarrow i}(t) = 0.$$

Proof. Indeed, since the edge (i, j) appears at least once in every period $[kB + 1, (k+1)B]$, it follows it must appear at least thrice within the period $[t+1, t+3B]$. If both $q_{i \rightarrow j}$ and $q_{j \rightarrow i}$ are zero during the first of those times, we are finished. If not, Lemma 3.11 implies they will be zero during either the second or the third of these times. \square

Our next lemma forms the core of the proof our main theorem. Informally, it states that if the variance is not too low, it must satisfy a quantitative decrease bound. Before proceeding, we need to mention the standard relationship between $W(t)$ and $V_2(t)$ which we will use:

$$\frac{W(t)}{2} \leq V_2(t) \leq \sqrt{n}W(t). \quad (3.10)$$

LEMMA 3.13. *Suppose that*

$$t \geq 2^{\frac{2}{1-\alpha}} \lceil 4(1+B)W(0) + 2B \rceil^{\frac{1}{\beta-\alpha}} + (20B\|x(0)\|_\infty)^{\frac{2}{1-\alpha}} + 20B$$

and

$$V_2(x(t)) \geq \frac{8n^{1.5}}{t^\alpha}. \quad (3.11)$$

Then

$$V_2(x(t+3B)) \leq (1 - \frac{1}{32n^4(t+3B)^\beta})V_2(x(t))$$

Proof. From Eq. (3.11), applying Eq. (3.10) we have,

$$W(t) > 8\frac{n}{t^\alpha}.$$

Because the subgraph consisting of core edges is connected, this implies that there exist two nodes i, j with the edge (i, j) being a core edge such that

$$|x_i(t) - x_j(t)| \geq \frac{W(x(t))}{n} \geq \frac{V_2(x(t))}{n^{1.5}} \quad (3.12)$$

It is an immediate consequence of the first inequality in this chain that

$$|x_i(t) - x_j(t)| > \frac{8}{t^\alpha} > \frac{4}{t^\alpha}. \quad (3.13)$$

Now our lower bound $t \geq 2\lceil 4(1+B)W(0) + 2B \rceil^{\frac{1}{\beta-\alpha}} + (20B\|x(0)\|_\infty)^{\frac{2}{1-\alpha}} + 20B$ allows us to apply Corollary 3.12 and claim there exists a time t' in $t' \in [t+1, t+3B]$ such that $q_{i \rightarrow j}(t') = q_{j \rightarrow i}(t') = 0$. Moreover, since between times t and $t' - 1$, both $x_i(t)$ and $x_j(t)$ can change by at most $3B(1/2)W(0)/t^\beta \leq (1/2)/t^\alpha$. This means $|x_i(t' - 1) - x_j(t' - 1)| > 7/t^\alpha > 7/(t')^\alpha$ and so

$$|\hat{x}_{i,j,\text{in}}(t') - \hat{x}_{i,j,\text{out}}(t')| > \frac{7}{(t')^\alpha} - 2\frac{1}{(t' - 2B - 1)^\alpha} > \frac{4}{(t')^\alpha},$$

where we used $t \geq 20B$ for the final inequality. The above equation implies that $j \in S(i, t')$ and consequently $i \in S(j, t')$.

We now proceed to lower bound the decrease from $V_2(x(t' - 1))$ to $V_2(x(t'))$ due to the link between i and j at time t' .

Since $j \in S(i, t')$ we have that $a_{ij}(t' - 1)$ is positive, and therefore $a_{ij}(t' - 1) \geq 1/(8 \max(d_i(t'), d_j(t')))$. Using the decomposition (see [21], [25])

$$A^2(t' - 1) = I - \sum_{i < j} [A^2(t' - 1)]_{ij} (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^T,$$

where \mathbf{e}_i as usual means the i 'th basis vector, we have that

$$\begin{aligned} V_2(A(t' - 1)x(t' - 1)) &= (x(t' - 1) - \bar{x}(t' - 1)\mathbf{1})^T (A(t' - 1))^2 (x(t' - 1) - \bar{x}\mathbf{1}) \\ &= V_2(x(t' - 1)) - \sum_{i < j} [A^2(t' - 1)]_{ij} (x_i(t' - 1) - x_j(t' - 1))^2, \end{aligned}$$

and, since $A(t'-1)$ is diagonally dominant and satisfies $a_{ij}(t'-1) \geq (1/8) \max(d_i(t'), d_j(t'))$, we have that $[A^2(t'-1)]_{ij} \geq (1/(16 \max(d_i(t'), d_j(t')))$; moreover, using Eq. (3.12), we obtain

$$V_2(A(t'-1)x(t'-1))^2 \leq V_2(x(t'-1))^2 - \frac{1}{16 \max(d_i(t'), d_j(t'))} \frac{V(x(t'-1))^2}{n^3}$$

so

$$V_2(A(t'-1)x(t'-1)) \leq V_2(x(t'-1)) \sqrt{1 - \frac{1}{16n^4}} \leq V_2(x(t'-1)) \left(1 - \frac{1}{32n^4}\right) \quad (3.14)$$

and therefore

$$\begin{aligned} V_2(x(t')) &\leq \left(\left(1 - \frac{1}{t^\beta}\right) V_2(x(t'-1)) + \frac{1}{t^\beta} V_2(A(t'-1)x(t'-1)) \right) \\ &\leq \left(1 - \frac{1}{32n^4(t')^\beta}\right) V_2(x(t'-1)) \end{aligned}$$

and since $V_2(x(t))$ is nonincreasing in t (Lemma 3.8) and $t' \leq t + 3B$, this concludes the proof. \square

We are almost ready to begin the proof of our main Theorem 2.3. However, we first need the following lemma which will allow us to bound some of the expressions which will appear shortly.

LEMMA 3.14. *The function $f(k) = \frac{e^{(1/d)(k+f)^c}}{k^e}$ with $c \in (0, 1)$, $d, e > 0$ is nondecreasing over the range $k \in [t_0, +\infty)$ where t_0 is any number which satisfies $\frac{t_0}{(t_0+f)^{1-c}} \geq de/c$.*

Proof. We have

$$f'(k) = \frac{e^{(1/d)(k+f)^c} (c/d)(k+f)^{c-1} k^e - e k^{e-1} e^{(1/d)(k+f)^c}}{k^{2e}}$$

so $f'(k) \geq 0$ if

$$\frac{c}{d}(k+f)^{c-1} \geq \frac{e}{k}$$

or

$$\frac{k}{(k+f)^{1-c}} \geq \frac{de}{c}.$$

\square

We now proceed to the proof of our main result. Following the numerous lemmas which we have already established, the proof now follows by some straightforward analysis.

Proof. [Proof of Theorem 2.3] Let us use the shorthand

$$T = 2^{\frac{2}{1-\alpha}} [4(1+B)W(0) + 2B]^{\frac{1}{\beta-\alpha}} + (20BV_\infty(0))^{\frac{2}{1-\alpha}} + 20B + (256n^4B)^{\frac{1}{1-\beta}}$$

and let us suppose $t \geq T$. If Eq. (3.11) holds at times $t, t+3B, t+6B, \dots, t+3kB$, then we can apply Lemma 3.13 to get

$$V_2(x(t+3kB)) \leq V_2(x(t)) \left(1 - \frac{1}{32n^4(t+3B)^\beta}\right) \left(1 - \frac{1}{32n^4(t+6B)^\beta}\right) \cdots \left(1 - \frac{1}{32n^4(t+3kB)^\beta}\right)$$

which, using the inequality $\ln(1-x) \leq -x$, implies

$$\ln \frac{V_2(x(t+3kB))}{V_2(x(t))} \leq - \sum_{j=1}^k \frac{1}{32n^4(t+3jB)^\beta} \leq - \frac{(1/(3B))(t+3kB)^{1-\beta} - (t+3B)^{1-\beta}}{32(1-\beta)n^4}$$

which in turn implies

$$V_2(x(t+3kB)) \leq V_2(x(t)) e^{-((t+3kB)^{1-\beta} - (t+3B)^{1-\beta}) / (128(1-\beta)n^4B)}$$

Now recall that this bound is derived under the assumption that Eq. (3.11) holds at times $t, t+3B, \dots, t+3kB$. By considering the the last time after T when Eq. (3.11) holds, we get the unconditional upper bound

$$V_2(x(t)) \leq \max(V_2(0) e^{-((t-3B)^{1-\beta} - (T+3B)^{1-\beta}) / (128(1-\beta)n^4B)}, \max_{T \leq k \leq t-6B} \frac{8n^{1.5}}{k^\alpha} e^{-((t-3B)^{1-\beta} - (k+3B)^{1-\beta}) / (128(1-\beta)n^4B)}, \frac{8n^{1.5}}{(t-6B)^\alpha}) \quad (3.15)$$

Lets consider how long it takes the first term to fall below ϵ . The inequality

$$V_2(0) e^{-((t-3B)^{1-\beta} - (T+3B)^{1-\beta}) / (128(1-\beta)n^4B)} \leq \epsilon$$

is implied by

$$(t-3B)^{1-\beta} - (T+3B)^{1-\beta} \geq 128n^4B \log \frac{V_2(0)}{\epsilon}$$

or

$$t \geq 3B + \left((T+3B)^{1-\beta} + 128n^4B \log \frac{V_2(0)}{\epsilon} \right)^{\frac{1}{1-\beta}}$$

which, using the inequality $(a+b)^k \leq 2^{k-1}(a^k + b^k)$ for $k \geq 1$, is in turn implied by

$$t \geq 3B + 2^{\frac{1}{1-\beta}} \left(T + 3B + (128n^4B \log \frac{\epsilon}{V(0)})^{\frac{1}{1-\beta}} \right)$$

Next, we consider the second maximum in our unconditional upper bound of Eq. (3.15). According to Lemma 3.14, if T is large enough so that

$$\frac{T}{(T+3B)^\beta} \geq 128n^4B\alpha$$

we may set $k = t - 6B$ in the maximization problem. But since $T \geq 20B$, we have that for this to be true it suffices that $T \geq (1.15 \cdot 128n^4B\alpha)^{1/(1-\beta)}$ which clearly holds. As a consequence, we finally have $V_2(x(t))$ falls below ϵ after

$$6B + \max \left(2^{\frac{1}{1-\beta}} \left(T + 3B + (128n^4B \log \frac{\epsilon}{V(0)})^{\frac{1}{1-\beta}} \right), \left(\frac{8n^{1.5}}{\epsilon} \right)^{\frac{1}{\alpha}} \right)$$

iterations. This is exactly the bound of Theorem 2.3 after a combination of terms.

□

REMARK 3.15. We now substantiate the claim, made earlier in the introduction, that the convergence result

$$\lim_{t \rightarrow \infty} x_i(t) = \frac{1}{n} \sum_{j=1}^n x_j(0)$$

holds for all B -connected sequences. Indeed, our proof here proceeds by arguing that, after some transient period, the following event occurs every $O(B)$ steps: there is a link (i, j) at time t with both messages $q_{i \rightarrow j}(t)$ and $q_{j \rightarrow i}(t)$ equal to zero, such that the “gap” $|x_i(t) - x_j(t)|$ is big (i.e., $|x_i(t) - x_j(t)| \geq V_2(x(t-3B))$). To get this result, we needed to assume the existence of a connected set of core edges which occur in every $[kB + 1, (k + 1)B]$.

If such a core set of edges does not exist, we still get the same conclusion but without an $O(B)$ bound on how long it takes for it to occur. This is sufficient to yield the convergence we claim, but not the convergence time bound in Theorem 2.3. Indeed, let us assume without loss of generality that $x_1(t) \leq x_2(t) \leq \dots \leq x_n(t)$ and let us consider the largest “gap” $|x_{i+1}(t) - x_i(t)|$. It is easy to see that there exists an i such that $|x_{i+1}(t) - x_i(t)| \geq V_2(x(t))/n^{1.5}$. What we need to show is that there is a later time $t' \geq t$ such that a node $i' \leq i$ and a node $j' \geq i + 1$ have the property that $q_{i' \rightarrow j'}(t') = q_{j' \rightarrow i'}(t') = 0$.

This follows by induction. The claim is trivial for $n = 1$. If it never occurs, the network effectively becomes disconnected into two strictly smaller networks, each achieving consensus by the inductive hypothesis. In that case, it is easy to see that $q_{i \rightarrow j}$ and $q_{j \rightarrow i}$ eventually equal zero for all edges (i, j) which occur infinitely often.

4. Simulations. Here we report on some simulations of our consensus protocol. We will be focusing on seeing how the performance of our protocol scales with the time t for fixed n and seeing how the time to reach a certain level of accuracy scales with the number of nodes for various graph topologies.

We will be simulating a slightly modified form of our protocol in which we omit the stepsize of $1/t^\beta$, i.e., in which we set $\beta = 0$. The introduction of this stepsize is a mathematical necessity for our convergence result, but it does not appear practically necessary. Similarly, we will double the weight every agent places on its neighbors as our consensus protocol is likely too conservative. Thus we will be simulating the update

$$x_i(t) = x_i(t-1) + \sum_{j \in S(i,t)} \frac{\hat{x}_{i,j,\text{in}}(t) - \hat{x}_{i,j,\text{out}}(t)}{2\max(d_i(t), d_j(t))}$$

instead of Eq. (2.1).

In Figure 4.1, we show the decay of the error $\max_i |x_i(t) - \frac{1}{n} \sum_{j=1}^n x_j(0)|$ with time for the complete graph and the line graph with $\alpha = 0.9$. Much like our error bound predict, we see a slow decay which appears to be asymptotic to a power of t . We point out the jagged features of the graph, which are more prominent in the case of the line: it often takes some time for the estimates $\hat{x}_{i,j,\text{in}}$ to become accurate, which results in periods without any updates.

In the next Figure 4.2, we are concerned with the time it takes our consensus protocol to achieve a certain level of error; we plot the time until $\max_i |x_i(t) -$

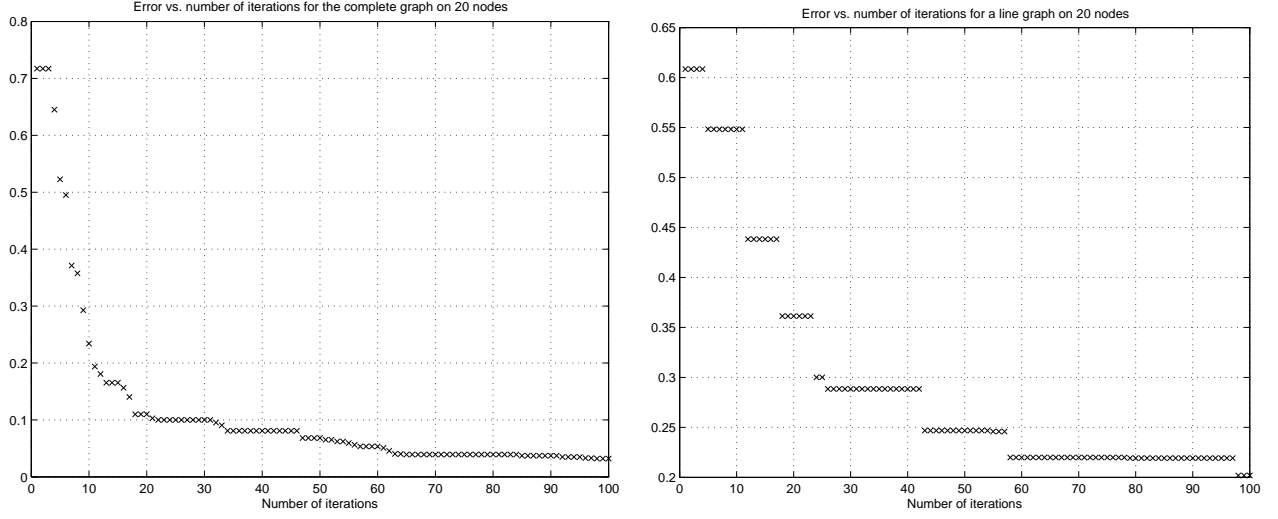


FIG. 4.1. The left plot shows our consensus protocol on the complete graph on 20 nodes, while the right graph shows our consensus protocol on the line graph on 20 nodes. Both graphs show the error at time t , i.e., $\max_i |x_i(t) - \frac{1}{n} \sum_{j=1}^n x_j(0)|$ vs the number of iterations t . Initial conditions are random for both graphs, and we have chosen $\alpha = 0.9$ in both.

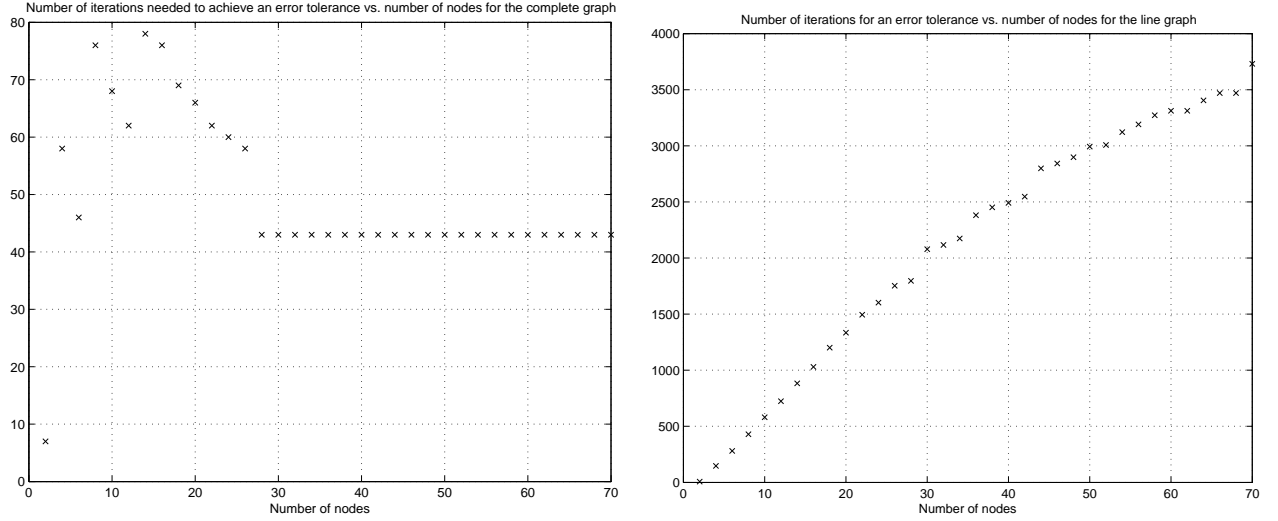


FIG. 4.2. Both graphs show the time it takes for our protocol to reach an error of 0.05, i.e., the time until $\max_i |x_i(t) - \frac{1}{n} \sum_{j=1}^n x_j(0)| \leq 0.05$. The number of nodes is shown on the x -axis. The left picture is for the complete graph while the right picture is for the line graph. The initial condition is the vector with $x_1(0) = 1$ and $x_k(0) = 1$ for $k = 2, \dots, n$ in both cases, and as before, $\alpha = 0.9$.

$\frac{1}{n} \sum_{j=1}^n x_j(0) \leq 0.05$ for the complete graph and the line graph. As expected, for the complete graph, the time does not grow with n . We note how irregular this time is for the complete graph. Indeed, the time it takes to attain a certain level of accuracy often appears discontinuous for our protocol, since our protocol itself is also discontinuous. Finally, we see a reasonably slow growth with n for the case of the

line.

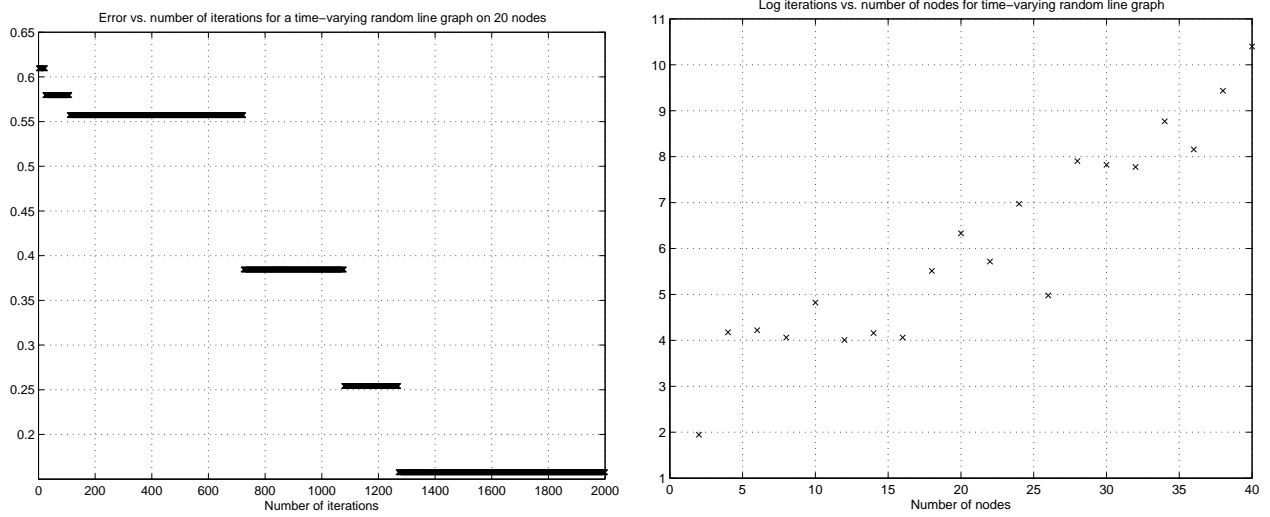


FIG. 4.3. Both graphs show our protocol on a time-varying line graph which is randomly generated at each stage. The left graph shows the decay of $\max_i |x_i(t) - \frac{1}{n} \sum_{j=1}^n x_j(0)|$ with time while the right graph shows the **logarithm** of the time it takes to achieve $\max_i |x_i(t) - \frac{1}{n} \sum_{j=1}^n x_j(0)| \leq 0.05$. The initial condition is random on the left and equal to the vector $x_1(0) = 1$ and $x_k(0) = 1$ for $k = 2, \dots, n$ on the right. In both cases, as before, $\alpha = 0.9$.

Finally, in Figure 4.3, we show our protocol on a time-varying graphs. We generate a random line at every stage by first relabeling the vertices uniformly at random and then connecting them in a standard line (i.e., putting an edge between 1 and 2, 2 and 3, and so forth). We see that the protocol appears to work, but the convergence time is considerably slower than for either of our fixed graph examples.

5. Conclusions. We have provided a consensus protocol in which nodes exchange ternary messages at every step. In contrast to the previous literature, our protocol can handle time-varying graph sequences, requires the exchange of only finitely many bits at every step, comes with a polynomial-time convergence rate, and does not need knowledge of either the sequence $G(t)$ or the number of agents to be implemented.

Our paper naturally raises a number of open questions. First, our protocol uses more storage than the standard consensus algorithm due to the need by each node i to maintain the estimates $\hat{x}_{i,j,\text{in}}, \hat{x}_{i,j,\text{out}}$. Is it possible to avoid this without using “unphysical” operations (for example, interleaving the digits of two real numbers to form a single real number)?

Secondly, our convergence time result was demonstrated to hold on B -core-connected graph sequences. However, most of the previous literature on consensus protocol has relied on the weaker assumption of B -connected graph sequences. An open question is to derive a similar convergence time result on the larger class of B -connected sequences.

Third, while the standard consensus protocol can be said to rely on broadcasts in the sense that each neighbor of node j measures or receives the same value $x_j(t)$ at time t , our protocol relies on individualized messages from every node to each of

its neighbors. We wonder whether there exists a consensus protocol with the same properties as ours (finitely many bits exchanged at each step, polynomial convergence rate on time-varying graphs, does not require knowledge of the graph sequence or of the number of which agents) which relies only on broadcasts by each node at every step.

Finally, the relation between speed, storage, and communication overhead in consensus is still poorly understood. There appear to be some tradeoffs involved between these quantities; for example, the current paper which puts aside the assumption that nodes can transmit real numbers derives a slower convergence time bound in contrast to the best average consensus convergence time [21] and has a larger amount of storage at each node. However, whether there are indeed any formal tradeoffs between these quantities is not clear at present.

REFERENCES

- [1] S. Boyd, A. Ghosh, B. Prabhakar, D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508-2530, 2006.
- [2] V.D. Blondel, J.M. Hendrickx, A. Olshevsky, J.N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," *Proceedings of the 44th IEEE Conference on Decision and Control*, Seville, Spain, 2005.
- [3] C. Gao, J. Cortes, F. Bullo, "Notes on averaging over acyclic graphs and discrete coverage control," *Automatica*, vol. 44, no. 8, pp. 2120-2127, 2008.
- [4] R. Carli, F. Bullo and S. Zampieri, "Quantized average consensus via dynamic coding/decoding schemes," *International Journal of Nonlinear and Robust Control*, vol. 20, no. 2, pp. 156-175, 2010.
- [5] R. Carli, A. Chiuso, L. Schenato, S. Zampieri, "Optimal synchronization for networks of noisy double integrators," *IEEE Transactions on Automatic Control*, vol. 56, no. 5, pp. 1146-1152, 2011.
- [6] R. Carli, F. Fagnani, P. Frasca and S. Zampieri, "Gossip consensus algorithms via quantized communication," *Automatica*, vol. 46, no. 1, pp. 70-80, 2010.
- [7] J.S. Caughman, G. Lafferriere, J.J.P. Veerman, and A. Williams, "Decentralized control of vehicle formations," *Systems and Control Letters*, vol. 54, no. 9, pp. 899-910, 2005.
- [8] H.L. Choi, L. Brunet, J.P. How, "Consensus-based decentralized auctions for robust task allocation," *IEEE Transactions on Robotics*, vol. 25, no. 4, pp. 912-926, 2009.
- [9] A. Dimakis, A.D. Sarwate, M. J. Wainwright, "Geographic gossip: efficient averaging for sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1205-1216, 2008.
- [10] J.A. Fax and R.M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1465-1476, 2004.
- [11] P. Frasca, R. Carli, F. Fagnani and S. Zampieri, "Average consensus on networks with quantized communication," *International Journal of Nonlinear and Robust Control*, vol. 19, no. 16, pp. 1787-1816, 2009.
- [12] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 3, pp. 988-1001, 2003.
- [13] A. Kashyap, T. Basar and R. Srikant, "Quantized consensus," *Automatica*, vol. 43, no. 7, pp. 1192-1203, 2007.
- [14] J.R. Lawton, R.W. Beard, and B. Young, "A decentralized approach to formation maneuvers," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 6, pp. 933-941, 2003.
- [15] T. Li, M. Fu, L. Xie and J. F. Zhang, "Distributed consensus with limited communication data rate," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 279-292, 2011.
- [16] J. Lavaei, R.M. Murray, "Quantized consensus by means of gossip algorithm," *IEEE Transactions on Automatic Control*, vol. 57, no. 1, pp. 19-32, 2012.
- [17] Q. Li, D. Rus, "Global clock synchronization in sensor networks," *IEEE Transactions on Computers*, vol. 55, no. 2, pp. 214-226, 2006.
- [18] S. Martinez, T. Karatas, F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243-255, 2004.
- [19] S. Martinez, J. Cortes, F. Bullo, "Motion coordination with distributed information," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 4, pp. 75-88, 2007.

- [20] L. Moreau, "Stability of multi-agent systems with time-dependent communication links," *IEEE Transactions on Automatic Control*, vol. 50, pp. 169-182, 2005.
- [21] A. Nedic, A. Olshevsky, A. Ozdaglar and J. N. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2506-2517, 2009.
- [22] R. Olfati-Saber, R.M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 3, pp. 1250-1533, 2004.
- [23] L. Schenato and F. Fiorentin, "Average timesynch: a consensus-based protocol for clock synchronization in wireless sensor networks," *Automatica*, vol. 47, no. 9, pp. 1878-1886, 2011.
- [24] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, no. 9, 1986, pp. 803-812.
- [25] L. Xiao, S. Boyd, S.J. Kim, "Distributed average consensus with least-mean square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33-46, 2007.